Solutions and improved perturbation analysis for the matrix equation $X - A^*X^{-p}A = Q$ $(p > 0)$

Jing Li

School of Mathematics and Statistics, Shandong University at Weihai, Weihai 264209, P.R. China

Abstract

In this paper the nonlinear matrix equation $X - A^*X^{-p}A = Q$ with $p > 0$ is investigated. We consider two cases of this equation: the case $p > 1$ and the case $0 < p < 1$. In the case $p > 1$, a new sufficient condition for the existence of a unique positive definite solution for the matrix equation is obtained. A perturbation estimate for the positive definite solution is derived. Explicit expressions of the condition number for the positive definite solution are given. In the case $0 < p < 1$, a new sharper perturbation bound for the unique positive definite solution is evaluated. A new backward error of an approximate solution to the unique positive definite solution is obtained. The theoretical results are illustrated by numerical examples.

Keywords: matrix equation, positive definite solution, perturbation bound, backward error, condition number

AMS classification: 15A24; 65H05

1. Introduction

In this paper we consider the Hermitian positive definite solution of the nonlinear matrix equation

$$X - A^*X^{-p}A = Q, \tag{1.1}$$

where A, Q and X are $n \times n$ complex matrices, Q is a positive definite matrix and $p > 0$. This type of nonlinear matrix equations arises in the analysis of ladder networks, the dynamic programming, control theory, stochastic filtering, statistics and many applications [1–4, 26, 27, 36].

In the last few years, Eq.(1.1) was investigated in some special cases. For the nonlinear matrix equations $X - A^*X^{-1}A = Q$ [11, 14, 18, 19, 23], $X - A^*X^{-2}A = Q$ [22, 40], $X - A^*X^{-n}A = Q$ [16, 17] and $X^s - A^*X^{-t}A = Q$ [25], there were many contributions in the literature to the solvability, numerical solutions and perturbation analysis. In addition, the similar equations $X+A^*X^{-1}A = Q$ [9, 10, 12, 14, 18, 19, 29, 37, 38], $X+A^*X^{-2}A = Q$ [21, 22, 39], $X+A^*X^{-n}A = Q$ [15, 17], $X^s+A^*X^{-t}A = Q$ [5, 6, 25, 34, 41], $X+A^*X^{-q}A = Q$ [13, 30, 35] and $X \pm \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q$ [7, 8, 20] were studied by many scholars.

☆The work was supported in part by Natural Science Foundation of China (11201263) and Natural Science Foundation of Shandong Province (ZR2012AQ004).

Email address: `xlijing@sdu.edu.cn` (Jing Li)

In [13], a sufficient condition for the equation $X - A^*X^{-p}A = Q$ $(0 < p \leq 1)$ to have a unique positive definite solution was provided. When the coefficient matrix A is nonsingular, several sufficient conditions for the equation $X - A^*X^{-q}A = Q$ $(q \geq 1)$ to have a unique positive definite solution were given in [33]. When the coefficient matrix A is an arbitrary complex matrix, necessary conditions and sufficient conditions for the existence of positive definite solutions for the equation $X-A^*X^{-q}A = Q$ $(q \geq 1)$ were derived in [31]. Li and Zhang in [24] proved that there always exists a unique positive definite solution to the equation $X - A^*X^{-p}A = Q$ $(0 < p < 1)$. They also obtained a perturbation bound and a backward error of an approximate solution for the unique solution of the equation $X - A^*X^{-p}A = Q$ $(0 < p < 1)$.

As a continuation of the previous results, the rest of the paper is organized as follows. Section 2 gives some preliminary lemmas that will be needed to develop this work. In Section 3, a new sufficient condition for Eq.(1.1) with $p > 1$ existing a unique positive definite solution is derived. In Section 4, a perturbation bound for the positive definite solution to Eq.(1.1) with $p > 1$ is given. In Section 5, applying the integral representation of matrix function, we also discuss the explicit expressions of condition number for the positive definite solution to Eq.(1.1) with $p > 1$. Furthermore, in Section 6, a new sharper perturbation bound for the unique positive definite solution to Eq.(1.1) with $0 < p < 1$ is evaluated. In Section 7, a new backward error of an approximate solution to Eq.(1.1) with $0 < p < 1$ is obtained. Finally, several numerical examples are presented in Section 8.

We denote by $C^{n\times n}$ the set of $n \times n$ complex matrices, by $\mathcal{H}^{n\times n}$ the set of $n \times n$ Hermitian matrices, by I the identity matrix, by $\|\cdot\|$ the spectral norm, by $\|\cdot\|_F$ the Frobenius norm and by $\lambda_{\max}(M)$ and $\lambda_{\min}(M)$ the maximal and minimal eigenvalues of M, respectively. For $A = (a_1, \ldots, a_n) = (a_{ij}) \in C^{n\times n}$ and a matrix B, $A \otimes B = (a_{ij}B)$ is a Kronecker product, and vec$A$ is a vector defined by vec$A = (a_1^T, \ldots, a_n^T)^T$. For $X, Y \in \mathcal{H}^{n\times n}$, we write $X \geq Y$(resp. $X > Y$) if $X - Y$ is Hermitian positive semi-definite (resp. definite). Let $\overline{\kappa} = \lambda_{\max}(A^*A)$, $\underline{\kappa} = \lambda_{\min}(A^*A)$.

2. Priliminaries

In this section, we will give some preliminary lemmas that will be needed to develop this work.

Lemma 2.1. *[24] For every positive definite matrix $X \in \mathcal{H}^{n\times n}$, if $0 < p < 1$, then*

(i) $X^{-p} = \dfrac{\sin p\pi}{\pi}\displaystyle\int_0^\infty (\lambda I + X)^{-1}\lambda^{-p}\, \mathrm{d}\lambda.$

(ii) $X^{-p} = \dfrac{\sin p\pi}{p\pi}\displaystyle\int_0^\infty (\lambda I + X)^{-1}X(\lambda I + X)^{-1}\lambda^{-p}\, \mathrm{d}\lambda.$

Lemma 2.2. *[24] There exists a unique positive definite solution X of $X - A^*X^{-p}A = Q$ $(0 < p < 1)$ and the iteration*

$$X_0 > 0, \quad X_n = Q + A^*X_{n-1}^{-p}A, \quad n = 1, 2, \cdots \tag{2.1}$$

converges to X.

Lemma 2.3. *[30]*

(i) *If $X \in \mathcal{H}^{n\times n}$, then $\|e^{-X}\| = e^{-\lambda_{\min}(X)}$.*

(ii) *If $X \in \mathcal{H}^{n\times n}$ and $r > 0$, then $X^{-r} = \frac{1}{\Gamma(r)}\int_0^\infty e^{-sX}s^{r-1}ds$.*

(ii) *If $A, B \in C^{n\times n}$, Then $e^{A+B} - e^A = \int_0^1 e^{(1-t)A}Be^{t(A+B)}dt$.*

3. A sufficient condition for the existence of a unique solution of $X - A^*X^{-p}A = Q$ $(p > 1)$

In this section, we derive a new sufficient condition for the existence of a unique solution of $X - A^*X^{-p}A = Q$ $(p > 1)$ beginning with the lemma.

Lemma 3.1. *[31]If*

$$\beta > (p\overline{\kappa})^{\frac{1}{p+1}}, \tag{3.1}$$

then Eq.(1.1) has a unique positive definite solution $X \in [\beta I,\ \alpha I]$, where α and β are respectively positive solutions of the following equations

$$(x - \lambda_{max}(Q))\left(\lambda_{min}(Q) + \frac{\underline{\kappa}}{x^p}\right)^p = \overline{\kappa}$$

and

$$(x - \lambda_{min}(Q))\left(\lambda_{max}(Q) + \frac{\overline{\kappa}}{x^p}\right)^p = \underline{\kappa}.$$

Furthermore,

$$\lambda_{min}(Q) \le \beta \le \alpha. \tag{3.2}$$

Theorem 3.2. *If*

$$((p\overline{\kappa})^{\frac{1}{p+1}} - \lambda_{min}(Q))\left(\lambda_{max}(Q) + \frac{\overline{\kappa}}{(p\overline{\kappa})^{\frac{p}{p+1}}}\right)^p < \underline{\kappa} \le \overline{\kappa} < \frac{\lambda_{max}(Q)\,(\lambda_{min}(Q)p)^p}{(p-1)^{p+1}}, \tag{3.3}$$

then Eq.(1.1) has a unique positive definite solution.

Proof. We first prove

$$\beta > (p\overline{\kappa})^{\frac{1}{p+1}}.$$

Let

$$f(x) = (x - \lambda_{min}(Q))\left(\lambda_{max}(Q) + \frac{\overline{\kappa}}{x^p}\right)^p - \underline{\kappa}.$$

By computaiton, we obtain

$$f'(x) = \frac{\overline{\kappa}}{x^p}\left(\lambda_{max}(Q) + \frac{\overline{\kappa}}{x^p}\right)^{p-1}\left(\frac{\lambda_{max}(Q)}{\overline{\kappa}}x^p + p^2\lambda_{min}(Q)x^{-1} + 1 - p^2\right).$$

Define that

$$g(x) = \frac{\lambda_{max}(Q)}{\overline{\kappa}}x^p + p^2\lambda_{min}(Q)x^{-1} + 1 - p^2.$$

Then $g(x)$ is decreasing on $[0,\ \left(\frac{\lambda_{min}(Q)p\overline{\kappa}}{\lambda_{max}(Q)}\right)^{\frac{1}{p+1}}]$ and increasing on $[\left(\frac{\lambda_{min}(Q)p\overline{\kappa}}{\lambda_{max}(Q)}\right)^{\frac{1}{p+1}},\ +\infty)$, which implies that

$$g_{\min} = g\left(\left(\frac{\lambda_{min}(Q)p\overline{\kappa}}{\lambda_{max}(Q)}\right)^{\frac{1}{p+1}}\right) = (1+p)\left(\frac{(\lambda_{min}(Q)p)^{\frac{p}{p+1}}\lambda_{max}^{\frac{1}{p+1}}(Q)}{(\overline{\kappa})^{\frac{1}{p+1}}} + 1 - p\right).$$

According to the condition $\overline{\kappa} < \frac{\lambda_{max}(Q)(\lambda_{min}(Q)p)^p}{(p-1)^{p+1}}$, it follows that $g_{\min} > 0$. Noting that

$$f'(x) = \frac{\overline{\kappa}}{x^p}\left(\lambda_{max}(Q) + \frac{\overline{\kappa}}{x^p}\right)^{p-1} g(x),$$

which implies that $f(x)$ is increasing on $(0, +\infty)$. Considering the condition (3.3), one sees that $f((p\overline{\kappa})^{\frac{1}{p+1}}) < 0$. Combining that and the definition of β in Lemma 3.1, we obtain $\beta > (p\overline{\kappa})^{\frac{1}{p+1}}$. By Lemma 3.1, Eq.(1.1) has a unique positive definite solution. $\square$

4. Perturbation bound for $X - A^*X^{-p}A = Q$ $(p > 1)$

Li and Zhang in [24] proved that there always exists a unique positive definite solution to the equation $X - A^*X^{-p}A = Q$ $(0 < p < 1)$. They also obtained a perturbation bound for the unique solution. But their approaches will become invalid for the case of $p > 1$. Since the equation $X - A^*X^{-p}A = Q$ $(p > 1)$ does not always have a unique positive definite solution, there are two difficulties for perturbation analysis to the equation $X - A^*X^{-p}A = Q$ $(p > 1)$. One difficulty is how to find some reasonable restrictions on the coefficient matrices of perturbed equation ensuring this equation has a unique positive definite solution. The other difficulty is how to find an expression of ΔX which is easy to handle.

Assume that the coefficient matrix A is perturbed to $\widetilde{A} = \Delta A + A$. Let $\widetilde{X} = \Delta X + X$ with $\Delta X \in \mathcal{H}^{n\times n}$ satisfying the perturbed equation

$$\widetilde{X} - \widetilde{A}^*\widetilde{X}^{-p}\widetilde{A} = Q, \quad p > 1. \tag{4.1}$$

In the following, we derive a perturbation estimate for the positive definite solution to the matrix equation $X - A^*X^{-p}A = Q$ $(p > 1)$ beginning with the lemma.

Lemma 4.1. *[31] If*

$$p\|A\|^2 < \lambda_{\min}^{p+1}(Q),$$

then Eq.(1.1) has a unique positive definite solution X, where $X \geq \lambda_{\min}(Q)I$.

Theorem 4.2. *If*

$$\|A\| < \sqrt{\frac{\lambda_{\min}^{p+1}(Q)}{p}} \ \ \textit{and} \ \ \|\Delta A\| < \sqrt{\frac{\lambda_{\min}^{p+1}(Q)}{p}} - \|A\|, \tag{4.2}$$

then

$$X - A^*X^{-p}A = Q \ \ \textit{and} \ \ \widetilde{X} - \widetilde{A}^*\widetilde{X}^{-p}\widetilde{A} = Q$$

have unique positive definite solutions X and $\widetilde{X}$, respectively. Furthermore,

$$\frac{\|\widetilde{X} - X\|}{\|X\|} \leq \frac{(2\|A\| + \|\Delta A\|)}{\lambda_{\min}^{p+1}(Q) - p\|A\|^2}\|\Delta A\| \equiv \varrho.$$

Proof. By (4.2), it follows that $\|\widetilde{A}\| \leq \|A\| + \|\Delta A\| \leq \sqrt{\frac{\lambda_{\min}^{p+1}(Q)}{p}}$. According to Lemma 4.1, the condition (4.2) ensures that Eq.(1.1) and Eq.(4.1) have unique positive definite solutions X and $\widetilde{X}$, respectively. Furthermore, we obtain that

$$X \geq \lambda_{\min}(Q)I, \quad \widetilde{X} \geq \lambda_{\min}(Q)I. \tag{4.3}$$

Subtracting (4.1) from (1.1) gives

$$\Delta X = \widetilde{A}^*\widetilde{X}^{-p}\widetilde{A} - A^*X^{-p}A = A^*(\widetilde{X}^{-p} - X^{-p})A + \Delta A^*\widetilde{X}^{-p}A + \widetilde{A}^*\widetilde{X}^{-p}\Delta A. \tag{4.4}$$

By Lemma 2.3 and inequalities in (4.3), we have

$$\begin{aligned}
& \|\Delta X + A^* X^{-p} A - A^* \widetilde{X}^{-p} A\| \\
= & \|\Delta X + A^* \frac{1}{\Gamma(p)} \int_0^\infty (e^{-sX} - e^{-s\widetilde{X}}) s^{p-1} ds A\| \\
= & \|\Delta X + A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s\widetilde{X}} \Delta X e^{-tsX} dt s^p ds A\| \\
\geq & \|\Delta X\| - \frac{\|A\|^2 \|\Delta X\|}{\Gamma(p)} \int_0^\infty \int_0^1 \|e^{-(1-t)s\widetilde{X}}\| \|e^{-tsX}\| dt s^p ds \\
\geq & \|\Delta X\| - \frac{\|A\|^2 \|\Delta X\|}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s\lambda_{\min}(\widetilde{X})} e^{-ts\lambda_{\min}(X)} dt s^p ds \\
\geq & \|\Delta X\| - \frac{\|A\|^2 \|\Delta X\|}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s\lambda_{\min}(Q)} e^{-ts\lambda_{\min}(Q)} dt s^p ds \\
= & \|\Delta X\| - \frac{\|A\|^2 \|\Delta X\|}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-s\lambda_{\min}(Q)} dt s^p ds \\
= & \|\Delta X\| - \frac{\Gamma(p+1)}{\Gamma(p)} \cdot \frac{\|A\|^2 \|\Delta X\|}{\lambda_{\min}^{p+1}(Q)} \\
= & \frac{\lambda_{\min}^{p+1}(Q) - p\|A\|^2}{\lambda_{\min}^{p+1}(Q)} \|\Delta X\|. \qquad (4.5)
\end{aligned}$$

Noting (4.2), we have

$$\lambda_{\min}^{p+1}(Q) - p\|A\|^2 > 0.$$

Combining (4.4) and (4.5), one sees that

$$\begin{aligned}
\frac{\lambda_{\min}^{p+1}(Q) - p\|A\|^2}{\lambda_{\min}^{p+1}(Q)} \|\Delta X\| & \leq \|\Delta A^* \widetilde{X}^{-p} A + \widetilde{A}^* \widetilde{X}^{-p} \Delta A\| \leq (\|\Delta A\| + 2\|A\|) \|\Delta A\| \|\widetilde{X}^{-p}\| \\
& \leq (\|\Delta A\| + 2\|A\|) \|\Delta A\| \lambda_{\min}^{-p}(Q),
\end{aligned}$$

which implies that

$$\frac{\|\Delta X\|}{\|X\|} \leq \frac{(\|\Delta A\| + 2\|A\|)}{\lambda_{\min}^{p+1}(Q) - p\|A\|^2} \|\Delta A\|.$$

□

5. Condition number for $X - A^* X^{-p} A = Q$ $(p > 1)$

A condition number is a measurement of the sensitivity of the positive definite stabilizing solutions to small changes in the coefficient matrices. In this section, we apply the theory of condition number developed by Rice [28] to derive explicit expressions of the condition number for the matrix equation $X - A^* X^{-p} A = Q$ $(p > 1)$.

Here we consider the perturbed equation

$$\widetilde{X} - \widetilde{A}^* \widetilde{X}^{-p} \widetilde{A} = \widetilde{Q}, \quad p > 1, \tag{5.1}$$

where $\widetilde{A}$ and $\widetilde{Q}$ are small perturbations of A and Q in Eq.(1.1), respectively.

Suppose that $p\|A\|^2 < \lambda_{\min}^{p+1}(Q)$ and $p\|\widetilde{A}\|^2 < \lambda_{\min}^{p+1}(\widetilde{Q})$. According to Lemma 4.1, Eq.(1.1) and Eq.(5.1) have unique positive definite solutions X and $\widetilde{X}$, respectively. Let $\Delta X = \widetilde{X} - X$, $\Delta Q = \widetilde{Q} - Q$ and $\Delta A = \widetilde{A} - A$.

Subtracting (5.1) from (1.1) gives

$$\begin{aligned}
\Delta X &= \widetilde{A}^* \widetilde{X}^{-p} \widetilde{A} - A^* X^{-p} A + \Delta Q = A^*(\widetilde{X}^{-p} - X^{-p})A + \Delta A^* \widetilde{X}^{-p} A + \widetilde{A}^* \widetilde{X}^{-p} \Delta A + \Delta Q \\
&= -A^* \frac{1}{\Gamma(p)} \int_0^\infty (e^{-sX} - e^{-s\widetilde{X}}) s^{p-1} ds A + \Delta A^* \widetilde{X}^{-p} A + \widetilde{A}^* \widetilde{X}^{-p} \Delta A + \Delta Q \\
&= -A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s\widetilde{X}} (\widetilde{X} - X) e^{-tsX} dt s^p ds A + \Delta A^* \widetilde{X}^{-p} A + \widetilde{A}^* \widetilde{X}^{-p} \Delta A + \Delta Q \\
&= -A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 (e^{-(1-t)s\widetilde{X}} - e^{-(1-t)sX}) \Delta X e^{-tsX} dt s^p ds A + \Delta Q \\
&\quad -A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)sX} \Delta X e^{-tsX} dt s^p ds A - (\widetilde{A}^* X^{-p} \Delta A - \widetilde{A}^* (X + \Delta X)^{-p} \Delta A) \\
&\quad + \widetilde{A}^* X^{-p} \Delta A - (\Delta A^* X^{-p} A - \Delta A^* (X + \Delta X)^{-p} A) + \Delta A^* X^{-p} A \\
&= A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 \int_0^1 e^{-(1-m)(1-t)sX} \Delta X e^{-m(1-t)s\widetilde{X}} \Delta X e^{-tsX} dm (1-t) dt s^{p+1} ds A + \Delta Q \\
&\quad -A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)sX} \Delta X e^{-tsX} dt s^p ds A + \Delta A^* X^{-p} \Delta A + A^* X^{-p} \Delta A + \Delta A^* X^{-p} A \\
&\quad -\widetilde{A}^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s(X+\Delta X)} \Delta X e^{-tsX} dt s^p ds \Delta A \\
&\quad -\Delta A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s(X+\Delta X)} \Delta X e^{-tsX} dt s^p ds A.
\end{aligned}$$

Therefore

$$\Delta X + A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)sX} \Delta X e^{-tsX} dt s^p ds A = E + h(\Delta X), \tag{5.2}$$

where

$$\begin{aligned}
&B = X^{-p} A, \\
&E = \Delta Q + (B^* \Delta A + \Delta A^* B) + \Delta A^* X^{-p} \Delta A, \\
&h(\Delta X) = A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 \int_0^1 e^{-(1-m)(1-t)sX} \Delta X e^{-m(1-t)s\widetilde{X}} \Delta X e^{-tsX} dm (1-t) dt s^{p+1} ds A \\
&\qquad - \widetilde{A}^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s(X+\Delta X)} \Delta X e^{-tsX} dt s^p ds \Delta A \\
&\qquad - \Delta A^* \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 e^{-(1-t)s(X+\Delta X)} \Delta X e^{-tsX} dt s^p ds A.
\end{aligned}$$

Lemma 5.1. *If*

$$p\|A\|^2 < \lambda_{\min}^{p+1}(Q), \tag{5.3}$$

then the linear operator $\mathbf{V} : \mathcal{H}^{n\times n} \to \mathcal{H}^{n\times n}$ *defined by*

$$\mathbf{V}W = W + \frac{1}{\Gamma(p)} \int_0^{\infty} \int_0^1 A^* e^{-(1-t)sX} W e^{-tsX} A dt s^p ds, \quad W \in \mathcal{H}^{n\times n}. \tag{5.4}$$

is invertible.

Proof. Define the operator $\mathbf{R} : \mathcal{H}^{n\times n} \to \mathcal{H}^{n\times n}$ by

$$\mathbf{R}Z = \frac{1}{\Gamma(p)} \int_0^{\infty} \int_0^1 A^* e^{-(1-t)sX} Z e^{-tsX} A dt s^p ds, \ Z \in \mathcal{H}^{n\times n};$$

it follows that

$$\mathbf{V}W = W + \mathbf{R}W.$$

Then $\mathbf{V}$ is invertible if and only if $I + \mathbf{R}$ is invertible.

According to Lemma 2.3 and the condition (5.3), we have

$$\begin{aligned}
\|\mathbf{R}W\| &\leq \|A\|^2\|W\|\frac{1}{\Gamma(p)}\int_0^{\infty}\int_0^1 \|e^{-(1-t)sX}\|\|e^{-tsX}\| dt s^p ds \\
&= \|A\|^2\|W\|\|\frac{1}{\Gamma(p)}\int_0^{\infty}\int_0^1 e^{-(1-t)s\lambda_{min}(X)} e^{-ts\lambda_{min}(X)} dt s^p ds \\
&\leq \|A\|^2\|W\|\|\frac{1}{\Gamma(p)}\int_0^{\infty}\int_0^1 e^{-(1-t)s\lambda_{min}(Q)} e^{-ts\lambda_{min}(Q)} dt s^p ds \\
&= \|A\|^2\|W\|\frac{1}{\Gamma(p)}\int_0^{\infty} e^{-s\lambda_{min}(Q)} s^p ds \\
&= \frac{p\|A\|^2}{\lambda_{min}^{p+1}(Q)}\|W\| < \|W\|,
\end{aligned}$$

which implies that $\|\mathbf{R}\| < 1$ and $I + \mathbf{R}$ is invertible. Therefore, the operator $\mathbf{V}$ is invertible. □

Thus, we can rewrite (5.2) as

$$\Delta X = \mathbf{V}^{-1}\Delta Q + \mathbf{V}^{-1}(B^*\Delta A + \Delta A^* B) + \mathbf{V}^{-1}(\Delta A^* X^{-p} \Delta A) + \mathbf{V}^{-1}(h(\Delta X)).$$

Obviously,

$$\Delta X = \mathbf{V}^{-1}\Delta Q + \mathbf{V}^{-1}(B^*\Delta A + \Delta A^* B) + O(\|(\Delta A,\ \Delta Q)\|_F^2), (\Delta A,\ \Delta Q) \to 0. \tag{5.5}$$

By the theory of condition number developed by Rice [6], we define the condition number of the Hermitian positive definite solution X to the matrix equation $X - A^*X^{-p}A = Q$ $(p > 1)$ by

$$c(X) = \lim_{\delta\to 0} \sup_{\|(\frac{\Delta A}{\eta}, \frac{\Delta Q}{\rho})\|_F \leq \delta} \frac{\|\Delta X\|_F}{\xi\delta}, \tag{5.6}$$

where ξ, η and ρ are positive parameters. Taking $\xi = \eta = \rho = 1$ in (5.6) gives the absolute condition number $c_{abs}(X)$, and taking $\xi = \|X\|_F$, $\eta = \|A\|_F$ and $\rho = \|Q\|_F$ in (5.6) gives the relative condition number $c_{rel}(X)$.

Substituting (5.5) into (5.6), we get

$$\begin{aligned} c(X) &= \frac{1}{\xi} \max_{\substack{(\frac{\Delta A}{\eta}, \frac{\Delta Q}{\rho}) \neq 0 \\ \Delta A \in \mathcal{C}^{n\times n}, \Delta Q \in \mathcal{H}^{n\times n}}} \frac{\|\mathbf{V}^{-1}(\Delta Q + B^*\Delta A + \Delta A^* B)\|_F}{\|(\frac{\Delta A}{\eta}, \frac{\Delta Q}{\rho})\|_F} \\ &= \frac{1}{\xi} \max_{\substack{(E,H) \neq 0 \\ E \in \mathcal{C}^{n\times n}, H \in \mathcal{H}^{n\times n}}} \frac{\|\mathbf{V}^{-1}(\rho H + \eta(B^*E + E^*B))\|_F}{\|(E,H)\|_F}. \end{aligned}$$

Let V be the matrix representation of the linear operator $\mathbf{V}$. Then it is easy to see that

$$V = I \otimes I + \frac{1}{\Gamma(p)} \int_0^\infty \int_0^1 (e^{-tsX}A)^T \otimes (A^* e^{-(1-t)sX}) dt s^p ds. \tag{5.7}$$

Let

$$\begin{aligned} & V^{-1} = S + i\Sigma, \\ & V^{-1}(I \otimes B^*) = V^{-1}(I \otimes (X^{-p}A)^*) = U_1 + i\Omega_1, \\ & V^{-1}(B^T \otimes I)\Pi = V^{-1}((X^{-p}A)^T \otimes I)\Pi = U_2 + i\Omega_2, \end{aligned} \tag{5.8}$$

$$S_c = \begin{bmatrix} S & -\Sigma \\ \Sigma & S \end{bmatrix}, \quad U_c = \begin{bmatrix} U_1 + U_2 & \Omega_2 - \Omega_1 \\ \Omega_1 + \Omega_2 & U_1 - U_2 \end{bmatrix}, \tag{5.9}$$

$$\mathrm{vec}H = x + \mathbf{i}y, \quad \mathrm{vec}E = a + \mathbf{i}b, \quad g = (x^T, y^T, a^T, b^T)^T,$$

where $x, y, a, b \in \mathcal{R}^{n^2}$, $S, \Sigma, U_1, U_2, \Omega_1, \Omega_2 \in \mathcal{R}^{n^2\times n^2}$, $M = (E, H)$, $\mathbf{i} = \sqrt{-1}$, Π is the vec-permutation matrix, i.e.,

$$\mathrm{vec}\, E^T = \Pi \,\mathrm{vec}\, E.$$

Furthermore, we obtain that

$$\begin{aligned} c(X) = \frac{1}{\xi} \max_{M \neq 0} & \frac{\|\mathbf{V}^{-1}(\rho H + \eta(B^*E + E^*B))\|_F}{\|(E,H)\|_F} \\ = \frac{1}{\xi} \max_{M \neq 0} & \frac{\|\rho V^{-1}\mathrm{vec}H + \eta V^{-1}((I \otimes B^*)\mathrm{vec}E + (B^T \otimes I)\mathrm{vec}E^*)\|}{\|(\mathrm{vec}E, \mathrm{vec}H)\|} \\ = \frac{1}{\xi} \max_{M \neq 0} & \frac{\|\rho(S + \mathbf{i}\Sigma)(x + \mathbf{i}y) + \eta[(U_1 + \mathbf{i}\Omega_1)(a + \mathbf{i}b) + (U_2 + \mathbf{i}\Omega_2)(a - \mathbf{i}b)]\|}{\|(\mathrm{vec}E, \mathrm{vec}H)\|} \\ = \frac{1}{\xi} \max_{g \neq 0} & \frac{\|(\rho S_c, \eta U_c)g\|}{\|g\|} \\ = \frac{1}{\xi} & \| (\rho S_c,\ \eta U_c)\|, \quad E \in \mathcal{C}^{n\times n}, H \in \mathcal{H}^{n\times n}. \end{aligned}$$

Then we have the following theorem.

Theorem 5.2. *If $p\|A\|^2 < \lambda_{\min}^{p+1}(Q)$, then the condition number $c(X)$ defined by (5.6) has the explicit expression*

$$c(X) = \frac{1}{\xi} \parallel (\rho S_c, \quad \eta U_c) \parallel, \tag{5.10}$$

where the matrices S_c and U_c are defined by (5.7)–(5.9).

Remark 1. From (5.10) we have the relative condition number

$$c_{rel}(X) = \frac{\parallel (\|Q\|_F S_c, \quad \|A\|_F U_c) \parallel}{\|X\|_F}. \tag{5.11}$$

5.1. The real case

In this subsection we consider the real case, i.e., all the coefficient matrices A, Q of the matrix equation $X - A^* X^{-p} A = Q$ $(p > 1)$ are real. In such a case the corresponding solution X is also real. Completely similar arguments as in Theorem 5.2 give the following theorem.

Theorem 5.3. *Let A, Q be real, $c(X)$ be the condition number defined by (5.6). If $p\|A\|^2 < \lambda_{\min}^{p+1}(Q)$, then $c(X)$ has the explicit expression*

$$c(X) = \frac{1}{\xi} \parallel (\rho S_r, \quad \eta U_r) \parallel,$$

where

$$\begin{aligned} S_r &= \left(I \otimes I + \frac{1}{\Gamma(p)} \int_0^{\infty} \int_0^1 (e^{-tsX} A)^T \otimes (A^* e^{-(1-t)sX}) dt s^p ds \right)^{-1}, \\ U_r &= S_r [I \otimes (A^T X^{-p}) + ((A^T X^{-p}) \otimes I)\Pi]. \end{aligned}$$

Remark 2. In the real case the relative condition number is given by

$$c_{rel}(X) = \frac{\parallel (\|Q\|_F S_r, \quad \|A\|_F U_r) \parallel}{\|X\|_F}.$$

6. New perturbation bound for $X - A^* X^{-p} A = Q$ $(0 < p < 1)$

Here we consider the perturbed equation

$$\widetilde{X} - \widetilde{A}^* \widetilde{X}^{-p} \widetilde{A} = \widetilde{Q}, \quad 0 < p < 1, \tag{6.1}$$

where $\widetilde{A}$ and $\widetilde{Q}$ are small perturbations of A and Q in Eq.(1.1), respectively. We assume that X and $\widetilde{X}$ are the solutions of Eq.(1.1) and Eq.(6.1), respectively. Let $\Delta X = \widetilde{X} - X$, $\Delta Q = \widetilde{Q} - Q$ and $\Delta A = \widetilde{A} - A$.

In this section we develop a new perturbation bound for the solution of Eq.(1.1) which is sharper than that in Theorem 3.1 [24].

Subtracting Eq.(1.1) from Eq.(6.1), using Lemma 2.1, we have

$$\Delta X + \frac{\sin p\pi}{\pi} \int_0^{\infty} [(\lambda I + X)^{-1} A]^* \Delta X [(\lambda I + X)^{-1} A] \lambda^{-p} d\lambda = E + h(\Delta X), \tag{6.2}$$

where

$$\begin{aligned}
B &= X^{-p}A,\\
E &= \Delta Q + (B^*\Delta A + \Delta A^* B) + \Delta A^* X^{-p}\Delta A, \qquad (6.3)\\
h(\Delta X) &= \frac{\sin p\pi}{\pi} A^* \int_0^{\infty} (\lambda I+X)^{-1}\Delta X(\lambda I+X+\Delta X)^{-1}\Delta X(\lambda I+X)^{-1}\lambda^{-p}\mathrm{d}\lambda A\\
&\quad - \frac{\sin p\pi}{\pi}\widetilde{A}^* \int_0^{\infty} (\lambda I+X)^{-1}\Delta X(\lambda I+X+\Delta X)^{-1}\lambda^{-p}\mathrm{d}\lambda\Delta A\\
&\quad - \frac{\sin p\pi}{\pi}\Delta A^* \int_0^{\infty} (\lambda I+X)^{-1}\Delta X(\lambda I+X+\Delta X)^{-1}\lambda^{-p}\mathrm{d}\lambda A.
\end{aligned}$$

By Lemma 5.1 in [24], the linear operator $\mathbf{L} : \mathcal{H}^{n\times n} \to \mathcal{H}^{n\times n}$ defined by

$$\mathbf{L}W = W + \frac{\sin p\pi}{\pi}\int_0^{\infty} [(\lambda I + X)^{-1}A]^* W[(\lambda I + X)^{-1}A]\lambda^{-p} d\lambda, \quad W \in \mathcal{H}^{n\times n}.$$

is invertible.

We also define operator $\mathbf{P} : \mathcal{C}^{n\times n} \to \mathcal{H}^{n\times n}$ by

$$\mathbf{P}Z = \mathbf{L}^{-1}(B^*Z + Z^*B), \quad Z \in \mathcal{C}^{n\times n}, \quad i = 1, 2, \cdots, m.$$

Thus,we can rewrite (6.2) as

$$\Delta X = \mathbf{L}^{-1}\Delta Q + \mathbf{P}\Delta A + \mathbf{L}^{-1}(\Delta A^* X^{-p}\Delta A) + \mathbf{L}^{-1}(h(\Delta X)). \qquad (6.4)$$

Define

$$\|\mathbf{L}^{-1}\| = \max_{\substack{W\in\mathcal{H}^{n\times n}\\ \|W\|=1}} \|\mathbf{L}^{-1}W\|, \qquad \|\mathbf{P}\| = \max_{\substack{Z\in\mathcal{C}^{n\times n}\\ \|Z\|=1}} \|\mathbf{P}Z\|.$$

Now we denote

$$\begin{aligned}
l &= \|\mathbf{L}^{-1}\|^{-1}, \quad \zeta = \|X^{-1}\|, \quad \xi = \|X^{-p}\|, \quad n = \|\mathbf{P}\|, \quad \eta = p\xi\|A\|^2\\
\epsilon &= \frac{1}{l}\|\Delta Q\| + n\|\Delta A\| + \frac{\xi}{l}\|\Delta A\|^2, \quad \sigma = \frac{p}{l}\zeta\xi(2\|A\| + \|\Delta A\|)\|\Delta A\|.
\end{aligned}$$

Theorem 6.1. *If*

$$\sigma < 1 \ \text{ and } \ \epsilon < \frac{l(1-\sigma)^2}{\zeta(l + l\sigma + 2\eta + 2\sqrt{(l\sigma+\eta)(\eta+l)})}, \qquad (6.5)$$

then

$$\|\widetilde{X} - X\| \le \frac{2l\epsilon}{l(1+\zeta\epsilon-\sigma) + \sqrt{l^2(1+\zeta\epsilon-\sigma)^2 - 4l\zeta\epsilon(l+\eta)}} \equiv \mu_*$$

Proof. Let

$$f(\Delta X) = \mathbf{L}^{-1}\Delta Q + \mathbf{P}\Delta A + \mathbf{L}^{-1}(\Delta A^* X^{-p}\Delta A) + \mathbf{L}^{-1}(h(\Delta X)).$$

Obviously, $f : \mathcal{H}^{n\times n} \to \mathcal{H}^{n\times n}$ is continuous. The condition (6.5) ensures that the quadratic equation $\zeta(l+\eta)x^2 - l(1+\zeta\epsilon-\sigma)x + l\epsilon = 0$ in x has two positive real roots. The smaller one is

$$\mu_* = \frac{2l\epsilon}{l(1+\zeta\epsilon-\sigma) + \sqrt{l^2(1+\zeta\epsilon-\sigma)^2 - 4l\zeta\epsilon(l+\eta)}}.$$

Define $\Omega = \{\Delta X \in \mathcal{H}^{n\times n} : \|\Delta X\| \le \mu_*\}$. Then for any $\Delta X \in \Omega$, by (6.5), we have

$$\begin{aligned}\|X^{-1}\Delta X\| &\le \|X^{-1}\|\|\Delta X\| \le \zeta\, \mu_* \le \zeta \cdot \frac{2\epsilon}{1+\epsilon-\sigma}\\ &= 1 + \frac{\zeta\epsilon+\sigma-1}{1+\zeta\epsilon-\sigma} \le 1 + \frac{-2(1-\sigma)(l\sigma+\eta)}{(l\sigma+l+2\eta)(1+\zeta\epsilon-\sigma)} < 1.\end{aligned}$$

It follows that $I - X^{-1}\Delta X$ is nonsingular and

$$\|I - X^{-1}\Delta X\| \le \frac{1}{1-\|X^{-1}\Delta X\|} \le \frac{1}{1-\zeta\|\Delta X\|}.$$

Therefore

$$\begin{aligned}\|f(\Delta X)\| &\le \frac{1}{l}\|\Delta Q\| + n\|\Delta A\| + \frac{\xi}{l}\|\Delta A_i\|^2 + \frac{p}{l}\zeta\xi\|A\|^2\frac{\|\Delta X\|^2}{1-\zeta\|\Delta X\|}\\ &+ \frac{p}{l}\zeta\xi(2\|A\| + \|\Delta A\|)\|\Delta A\| \cdot \frac{\|\Delta X\|}{1-\zeta\|\Delta X\|}\\ &\le \epsilon + \frac{\sigma\|\Delta X\|}{1-\zeta\|\Delta X\|} + \frac{\eta\zeta\|\Delta X\|^2}{l(1-\zeta\|\Delta X\|)}\\ &\le \epsilon + \frac{\sigma\mu_*}{1-\zeta\mu_*} + \frac{\theta\zeta\mu_*^2}{l(1-\zeta\mu_*)} = \mu_*,\end{aligned}$$

for $\Delta X \in \Omega$. That is $f(\Omega) \subseteq \Omega$. According to Schauder fixed point theorem, there exists $\Delta X_* \in \Omega$ such that $f(\Delta X_*) = \Delta X_*$. It follows that $X + \Delta X_*$ is a Hermitian solution of Eq.(6.1). By Lemma 2.2, we know that the solution of Eq.(6.1) is unique. Then $\Delta X_* = \widetilde{X} - X$ and $\|\widetilde{X} - X\| \le \mu_*$. □

7. New backward error for $X - A^*X^{-p}A = Q$ $(0 < p < 1)$

In this section we evaluate a new backward error of an approximate solution to the unique solution, which is sharper than that in Theorem 4.1 [24].

Theorem 7.1. *Let $\widetilde{X} > 0$ be an approximation to the solution X of (1.1). If $\|\widetilde{X}^{-\frac{p}{2}}A\|^2\|\widetilde{X}^{-1}\| < 1$ and the residual $R(\widetilde{X}) \equiv Q + A^*\widetilde{X}^{-p}A - \widetilde{X}$ satisfies*

$$\|R(\widetilde{X})\| \le \frac{\theta_1}{2}\min\left\{1, \frac{\theta_1}{2\lambda_{\min}(\widetilde{X})}\right\}, \text{ where } \theta_1 = (1 - \|\widetilde{X}^{-\frac{p}{2}}A\|^2\|\widetilde{X}^{-1}\|)\lambda_{\min}(\widetilde{X}) + \|R(\widetilde{X})\| > 0, \quad (7.1)$$

then

$$\|\widetilde{X} - X\| \le \theta\|R(\widetilde{X})\|, \text{ where } \theta = \frac{2\lambda_{\min}(\widetilde{X})}{\theta_1 + \sqrt{\theta_1^2 - 4\lambda_{\min}(\widetilde{X})\|R(\widetilde{X})\|}}. \quad (7.2)$$

To prove the above theorem, we first verify the following lemma.

Lemma 7.2. *For every positive definite matrix $X \in \mathcal{H}^{n\times n}$, $0 < p < 1$, if $X + \Delta X \ge (1/\nu)I > 0$, then*

$$\|A^*((X + \Delta X)^{-p} - X^{-p})A\| \le p\,(\|\Delta X\| + \nu\|\Delta X\|^2)\|X^{-\frac{p}{2}}A\|^2\|X^{-1}\|. \quad (7.3)$$

Proof. If $X + \Delta X \geq (1/\nu)I > 0$, then

$$\begin{aligned}
& \|A^*((X+\Delta X)^{-p} - X^{-p})A\| \\
= & \left\| A^* \left(\frac{\sin p\pi}{\pi} \int_0^\infty \left((\lambda I + X + \Delta X)^{-1} - (\lambda I + X)^{-1} \right) \lambda^{-p} \, \mathrm{d}\lambda \right) A \right\| \\
\leq & \frac{\sin p\pi}{\pi} \left(\|A^* \int_0^\infty (\lambda I + X)^{-1} \Delta X (\lambda I + X)^{-1} \lambda^{-p} \mathrm{d}\lambda \, A\| \right) \\
& + \frac{\sin p\pi}{\pi} \left(\|A^* \int_0^\infty (\lambda I + X)^{-1} \Delta X (\lambda I + X + \Delta X)^{-1} \Delta X (\lambda I + X)^{-1} \lambda^{-p} \mathrm{d}\lambda \, A\| \right) \\
\leq & p \, \|A^* X^{-p} A\| \|X^{-1}\| \|\Delta X\| + p \, \|A^* X^{-p} A\| \nu \|\Delta X\|^2 \|X^{-1}\| \\
= & p \, (\|\Delta X\| + \nu \|\Delta X\|^2) \|X^{-\frac{p}{2}} A\|^2 \|X^{-1}\|.
\end{aligned}$$

□

Proof. Let

$$\Psi = \{\Delta X \in \mathcal{H}^{n\times n} : \|\Delta X\| \leq \theta \|R(\widetilde{X})\|\}.$$

Obviously, Ψ is a nonempty bounded convex closed set. Let

$$g(\Delta X) = A^*((\widetilde{X} + \Delta X)^{-p} - \widetilde{X}^{-p})A + R(\widetilde{X}).$$

Evidently $g : \Psi \mapsto \mathcal{H}^{n\times n}$ is continuous. We will prove that $g(\Psi) \subseteq \Psi$. For every $\Delta X \in \Psi$, we have

$$\Delta X \geq -\theta \|R(\widetilde{X})\| I.$$

Hence

$$\widetilde{X} + \Delta X \geq \widetilde{X} - \theta \|R(\widetilde{X})\| I \geq (\lambda_{\min}(\widetilde{X}) - \theta \|R(\widetilde{X})\|) I.$$

Using (7.1) and (7.2), one sees that

$$\theta \|R(\widetilde{X})\| = \frac{2\lambda_{\min}(\widetilde{X}) \|R(\widetilde{X})\|}{\theta_1 + \sqrt{\theta_1^2 - 4\lambda_{\min}(\widetilde{X}) \|R(\widetilde{X})\|}} < \frac{2\lambda_{\min}(\widetilde{X}) \|R(\widetilde{X})\|}{\theta_1} < \lambda_{\min}(\widetilde{X}).$$

Therefore, $(\lambda_{\min}(\widetilde{X}) - \theta \|R(\widetilde{X})\|) I > 0$.

According to (7.3), we obtain

$$\begin{aligned}
& \|g(\Delta X)\| \\
\leq & p \, (\|\Delta X\| + \frac{\|\Delta X\|^2}{\lambda_{\min}(\widetilde{X}) - \theta \|R(\widetilde{X})\|}) \|X^{-\frac{p}{2}} A\|^2 \|\widetilde{X}^{-1}\| + \|R(\widetilde{X})\| \\
\leq & \left(\theta \|R(\widetilde{X})\| + \frac{(\theta \|R(\widetilde{X})\|)^2}{\lambda_{\min}(\widetilde{X}) - \theta \|R(\widetilde{X})\|} \right) (p \|X^{-\frac{p}{2}} A\|^2 \|\widetilde{X}^{-1}\|) + \|R(\widetilde{X})\| \\
= & \theta \|R(\widetilde{X})\|.
\end{aligned}$$

By Brouwer's fixed point theorem, there exists a $\Delta X \in \Psi$ such that $g(\Delta X) = \Delta X$. Hence $\widetilde{X} + \Delta X$ is a solution of Eq.(1.1). Moreover, by Lemma 2.2, we know that the solution X of Eq.(1.1) is unique. Then

$$\|\widetilde{X} - X\| = \|\Delta X\| \leq \theta \|R(\widetilde{X})\|.$$

□

8. Numerical Examples

To illustrate the theoretical results of the previous sections, in this section four simple examples are given, which were carried out using MATLAB 7.1. For the stopping criterion we take $\varepsilon_{k+1}(X) = \|X_k - A^* X_k^{-p} A - Q\| < 1.0e-10$.

Example 8.1. *We consider the matrix equation*

$$X - A^* X^{-\frac{1}{3}} A = I,$$

where

$$A = \frac{A_0}{\|A_0\|}, \quad A_0 = \begin{pmatrix} 2 & 0.95 \\ 0 & 1 \end{pmatrix}.$$

Suppose that the coefficient matrix A is perturbed to $\widetilde{A} = A + \Delta A$, where

$$\Delta A = \frac{10^{-j}}{\|C^T + C\|}(C^T + C)$$

*and C is a random matrix generated by MATLAB function **randn**.*

We compare our own result $\frac{\mu_*}{\|X\|}$ in Theorem 6.1with the perturbation bound ξ_* proposed in Theorem 3.1 [24].

The condition in Theorem 3.1 [24] is

$$con1 = \sqrt{\|A\|^2 + \zeta} - \|A\| - \|\Delta A\| > 0.$$

The conditions in Theorem 6.1 are

$$con2 = 1 - \sigma > 0, \quad con3 = \frac{l(1-\sigma)^2}{\zeta(l + \sigma l + 2\eta + 2\sqrt{(l\sigma + \eta)(\eta + l)})} - \epsilon > 0.$$

By computation, we list them in Table 1.

Table 1: Conditions for Example 8.1 with different values of j

j	4	5	6	7
$con1$	0.0455	0.0456	0.0456	0.0456
$con2$	0.9999	1.0000	1.0000	1.0000
$con3$	0.3957	0.3959	0.3959	0.3959

The results listed in Table 1 show that the conditions in Theorem 3.1 [24] and Theorem 6.1 are satisfied.

By Theorem 3.1 in [24] and Theorem 6.1, we can compute the relative perturbation bounds ξ_*, $\frac{\mu_*}{\|X\|}$, respectively. These results averaged as the geometric mean of 10 randomly perturbed runs. Some results are listed in Table 2.

The results listed in Table 2 show that the perturbation bound $\frac{\mu_*}{\|X\|}$ given by Theorem 6.1 is fairly sharp, while the bound ξ_* given by Theorem 3.1 in[24] is conservative.

Table 2: Results for Example 8.1 with different values of j

j	4	5	6	7
$\frac{\|\widetilde{X}-X\|}{\|X\|}$	6.8119×10^{-5}	4.2332×10^{-6}	4.3287×10^{-7}	5.5767×10^{-8}
ξ_*	2.6003×10^{-4}	2.1375×10^{-5}	1.9229×10^{-6}	2.7300×10^{-7}
$\frac{\mu_*}{\|X\|}$	8.8966×10^{-5}	6.5825×10^{-6}	7.2867×10^{-7}	9.3455×10^{-8}

Example 8.2. *Consider the equation*

$$X - A^*X^{-3/4}A = Q,$$

for

$$A = \begin{pmatrix} 0.2 & -0.2 \\ 0.1 & 0.1 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.8939 & 0.2987 \\ 0.1991 & 0.6614 \end{pmatrix}.$$

Choose $\widetilde{X}_0 = 3Q$. Let the approximate solution $\widetilde{X}_k$ be given with the iterative method (2.1), where k is the iterative number. Assume that the solution X of Eq.(1.1) is unknown.

We compare our own result with the backward error proposed in Theorem 4.1 [24].

The residual $R(\widetilde{X}_k) \equiv Q + A^*\widetilde{X}_k^{-p}A - \widetilde{X}_k$ satisfies the conditions in Theorem 4.1 [24] and in Theorem 7.1.

By Theorem 4.1 in[24] , we can compute the backward error bound

$$\| \widetilde{X}_k - X \| \leq \nu_* \|R(\widetilde{X}_k)\|, \quad where \quad \nu_* = \frac{2\|\widetilde{X}_k\|\|\widetilde{X}_k^{-1}\|}{1 - \frac{3}{4}\|\widetilde{X}_k^{-\frac{3}{8}} A \widetilde{X}_k^{-1/2}\|^2}.$$

By Theorem 7.1, we can compute the new backward error bound

$$\|\widetilde{X}_k - X\| \leq \theta \|R(\widetilde{X}_k)\|, \quad where \quad \theta = \frac{2\lambda_{\min}(\widetilde{X}_k)}{\theta_1 + \sqrt{\theta_1^2 - 4\lambda_{\min}(\widetilde{X}_k)\|R(\widetilde{X}_k)\|}},$$

$\theta_1 = (1 - \|\widetilde{X}_k^{-\frac{3}{8}} A\|^2 \|\widetilde{X}_k^{-1}\|)\lambda_{\min}(\widetilde{X}_k) + \|R(\widetilde{X}_k)\|.$

Let

$$\kappa_1 = \frac{\nu_* \|R(\widetilde{X}_k)\|}{\| \widetilde{X}_k - X \|}, \quad \kappa_2 = \frac{\theta \|R(\widetilde{X}_k)\|}{\| \widetilde{X}_k - X \|}.$$

Some results are shown in Table3.

From the results listed in Table 3 we see that the new backward error bound $\theta\|R(\widetilde{X}_k)\|$ is sharper than the backward error bound $\nu_*\|R(\widetilde{X}_k)\|$ in [24]. Moreover, we see that the backward error $\theta\|R(\widetilde{X})\|$ for an approximate solution $\widetilde{X}$ seems to be independent of the conditioning of the solution X.

Example 8.3. *We consider the matrix equation*

$$X - A^*X^{-3}A = 5I,$$

Table 3: Results for Example 8.2 with different values of k

k	4	5	6	7
$\|\widetilde{X}_k - X\|$	6.2131×10^{-6}	1.5830×10^{-7}	8.2486×10^{-9}	6.0132×10^{-10}
$\nu_*\|R(\widetilde{X}_k)\|$	2.5930×10^{-5}	6.6257×10^{-7}	3.5697×10^{-8}	2.4646×10^{-9}
κ_1	4.1734	4.1856	4.3277	4.0986
$\theta\|R(\widetilde{X}_k)\|$	7.0053×10^{-6}	1.7900×10^{-7}	9.6440×10^{-9}	6.6583×10^{-10}
κ_2	1.1275	1.1308	1.1692	1.1073

where

$$A = \frac{A_0}{\|A_0\|}, \quad A_0 = \begin{pmatrix} 2 & 0.95 \\ 0 & 1 \end{pmatrix}.$$

We now consider the perturbation bounds for the solution X when the coefficient matrix A is perturbed to $\widetilde{A} = A + \Delta A$, where

$$\Delta A = \frac{10^{-j}}{\|C^T + C\|}(C^T + C)$$

and C is a random matrix generated by MATLAB function ***randn***.

The conditions in Theorem 4.2 are satisfied.

By Theorem 4.2, we can compute the relative perturbation bound ϱ with different values of j. These results averaged as the geometric mean of 10 randomly perturbed runs. Some results are listed in Table 4.

Table 4: Results for Example 8.3 with different values of j

j	4	5	6	7
$\frac{\|\widetilde{X}-X\|}{\|X\|}$	1.1892×10^{-7}	2.1101×10^{-8}	2.4085×10^{-9}	1.6847×10^{-10}
ϱ	2.0791×10^{-7}	3.5353×10^{-8}	3.9573×10^{-9}	3.2580×10^{-10}

The results listed in Table 4 show that the perturbation bound ϱ given by Theorem 4.2 is fairly sharp.

Example 8.4. *Consider the matrix equation $X - A^*X^{-3}A = Q$, where*

$$A = \begin{pmatrix} 0.5 & 0.55 - 10^{-k} \\ 1 & 1 \end{pmatrix}, \qquad Q = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}.$$

By Remark 2, we can compute the relative condition number $c_{rel}(X)$. Some results are listed in Table 5.

The numerical results listed in the second line show that the unique positive definite solution X is well-conditioned.

Acknowledgements

The authors would like to express their gratitude to the referees for their fruitful comments and suggestions regarding the earlier version of this paper.

Table 5: Results for Example 8.4 with different values of k

k	1	3	5	7	9
$c_{rel}(X)$	1.2510	1.0991	1.0009	1.0009	1.0009

References

[1] W. N. Anderson, G. B. Kleindorfer, M. B. Kleindorfer, and M. B. Woodroofe, Consistent estimates of the parameters of a linear systems, Ann. Math. Statist. 40 (1969) 2064–2075.

[2] W. N. Anderson, T. D. Morley, and G. E. Trapp, The cascade limit, the shorted operator and quadratic optimal control, in Linear Circuits, Systems and Signal Processsing: Theory and Application (Christopher I. Byrnes, Clyde. F. Martin, and Richard E. Saeks, Eds.), North-Holland, New York, 1988, 3–7.

[3] R. S. Bucy, A priori bounds for the Riccati equation, in proceedings of the Berkley Symposium on Mathematical Statistics and Probability, Vol. III: Probability Theory, Univ. of California Press, Berkeley, 1972, 645–656.

[4] B. L. Buzbee, G. H. Golub, C. W. Nielson, On direct methods for solving Poisson's equations, SIAM J. Numer. Anal. 7 (1970)627-656.

[5] J. Cai, G.L. Chen, On the Hermitian positive definite solution of nonlinear matrix equation $X^s + A^*X^{-t}A = Q$, Appl. Math. Comput. 217 (2010) 2448–2456.

[6] X.F. Duan, A.P. Liao, on the existence of Hermitian positive definite solutions of the matrix equation $X^s + A^*X^{-t}A = Q$, Linear Algebra Appl. 429 (2008) 673–687.

[7] X.F. Duan, A.P. Liao, B.Tang, On the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$, Linear Algebra Appl. 429 (2008) 110-121.

[8] X.F. Duan, C.M. Li, A.P. Liao, Solutions and perturbation analysis for the nonlinear matrix equation $X + \sum_{i=1}^{m} A_i^* X^{-1} A_i = I$, Appl. Math. Comput. 218 (2011) 4458-4466.

[9] J.C. Engwerda, On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$, Linear Algebra Appl. 194 (1993) 91–108.

[10] J.C. Engwerda, A.C.M. Ran, A.L. Rijkeboer, Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = Q$, Linear Algebra Appl. 186 (1993) 255–275.

[11] A. Ferrante, Hermitian solutions of the equation $X = Q + NX^{-1}N^*$, Linear Algebra Appl. 247 (1996) 359–373.

[12] C.H. Guo, P. Lancaster, Iterative solution of two matrix equations, Math. Comp. 68 (1999) 1589–1603.

[13] V.I. Hasanov, Positive definite solutions of the matrix equations $X \pm A^*X^{-q}A = Q$, Linear Algebra Appl. 404 (2005) 166–182.

[14] V.I. Hasanov, Notes on two perturbation estimates of the extreme solutions to the equations $X \pm A^*X^{-1}A = Q$, Appl. Math. Comput. 216 (2010), 1355–1362.

[15] V.I. Hasanov, On positive definite solutions of the family of matrix equation $X + A^*X^{-n}A = Q$, J. Comput. Appl. Math. 193(2006), 277–301.

[16] V.I. Hasanov, I.G. Ivanov, On the matrix equation $X - A^*X^{-n}A = I$, Appl. Math. Comput. 168 (2005), 1340–1356.

[17] V.I. Hasanov, I.G. Ivanov, Solutions and perturbation estimates of the matrix equation $X \pm A^*X^{-n}A = Q$, Linear Algebra Appl. 156 (2004), 513–525.

[18] V.I. Hasanov, I.G. Ivanov, On two perturbation estimates of the extreme solutions to the equations $X \pm A^*X^{-1}A = Q$, Linear Algebra Appl. 413 (2006), 81–92.

[19] V.I. Hasanov, I.G. Ivanov, F. Uhlig, Improved perturbation estimates for the matrix equation $X \pm A^*X^{-1}A = Q$, Linear Algebra Appl. 379 (2004) 113–135.

[20] Y.M. He, J.H. Long, On the Hermitian positive definite solution of the nonlinear matrix equation $X + \sum_{i=1}^{m} A_i^* X^{-1} A_i = I$, Appl. Math. Comput. 216 (2010) 3480-3485.

[21] I.G. Ivanov, S.M. El-Sayed, Properties of positive definite solution of the equation $X + A^*X^{-2}A = I$, Linear Algebra Appl. 279(1998) 303–316.

[22] I.G. Ivanov, V.I. Hasanov, B.V. Minchev, On matrix equations $X \pm A^*X^{-2}A = I$, Linear Algebra Appl. 326 (2001) 27–44.

[23] J. Li, The Hermitian positive definite solutions and perturbation analysis of the matrix equation $X - A^*X^{-1}A = Q$, Math. Numer. Sinica, 30 (2008) 129–142. (in Chinese).

[24] J. Li, Y.H. Zhang, Perturbation analysis of the matrix equation $X - A^*X^{-q}A = Q$, Linear Algebra Appl. 431 (2009) 1489–1501.

[25] X.G. Liu, H. Gao, On the positive definite solutions of the matrix equations $X^s \pm A^T X^{-t} A = I_n$, Linear Algebra Appl. 368 (2003) 83–97.
[26] D. V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981) 187–295.
[27] W. Pusz and S. L. Woronowitz, Functional caculus for sequlinear forms and purification map, Rep. Math. Phys. 8 (1975)159–170.
[28] J.R. Rice, A theory of condition, SIAM J. Numer. Anal. 3 (1966) 287-310.
[29] J.G. Sun, S.F. Xu, perturbation analysis of the maximal solution of the matrix equation $X + A^* X^{-1} A = P$. II, Linear Algebra Appl. 362 (2003) 211–228.
[30] J.F. Wang, Y.H.Zhang, B.R. Zhu, The Hermitian positive definite solutions of the matrix equation $X + A^* X^{-q} A = I (q > 0)$, Math. Numer. Sinica, 26 (2004) 61–72. (in Chinese).
[31] H. Xiao, J.T. Wang, On the matrix equation $X - A^* X^{-p} A = Q$ $(p > 1)$, Chin. J. Eng. Math. 26 (2009) 305–309.
[32] S.F. Xu, Perturbation analysis of the maximal solution of the matrix equation $X + A^* X^{-1} A = P$, Linear Algebra Appl. 336 (2001) 61–70.
[33] X.Y. Yin, S.Y. Liu, Positive definite solutions of the matrix equations $X \pm A^* X^{-q} A = Q$ $(q \geq 1)$, Comput. Math. Appl. 59 (2010) 3727–3739.
[34] X.Y. Yin, S.Y. Liu, L. Fang, Solutions and perturbation estimates for the matrix equation $X^s + A^* X^{-t} A = Q$, Linear Algebra Appl. 431 (2009) 1409-1421.
[35] X.Y. Yin, S.Y. Liu, T.X. Li, ON positive definite solutions of the matrix equation $X + A^* X^{-q} A = Q$ $(0 < q \leq 1)$, TAIWAN J. Math. 16 (2012) 1391–1407.
[36] J. Zabezyk, Remarks on the control of discrete time distributed parameter systems, SIAM J. Control. 12 (1974) 721–735.
[37] X. Zhan, Computing the extremal positive definite solutions of a matrix equations, SIAM J. Sci. Comput. 17 (1996) 1167–1174.
[38] X. Zhan, J. Xie, On the matrix equation $X + A^* X^{-1} A = I$, Linear Algebra Appl. 247 (1996) 337–345.
[39] Y.H. Zhang, On Hermitian positive definite solutions of matrix equation $X + A^* X^{-2} A = I$, Linear Algebra Appl. 372 (2003) 295–304.
[40] Y.H. Zhang, On Hermitian positive definite solutions of matrix equation $X - A^* X^{-2} A = I$, J. Comp. Math. 23 (2005) 408–418.
[41] D.M. Zhou, G.L. Chen, G.X. Wu, X.Y. zhang, Some properties of the nonlinear matrix equation $X^s + A^* X^{-t} A = Q$, J. Math. Anal. Appl. 392 (2012) 75-82.